

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: Hodgson, John
Lawlor, Elizabeth

(ii) TITLE OF THE INVENTION: Novel tRNA Synthetase

(iii) NUMBER OF SEQUENCES: 2

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: SmithKline Beecham Corporation
(B) STREET: 709 Swedeland Road
(C) CITY: King of Prussia
(D) STATE: PA
(E) COUNTRY: USA
(F) ZIP: 19406-0939

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ for Windows Version 2.0

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE: 17-JAN-1997
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 9601095.4
(B) FILING DATE: 19-JAN-1996

(A) APPLICATION NUMBER: 9615845.6
(B) FILING DATE: 27-JUL-1996

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Gimmi, Edward R
(B) REGISTRATION NUMBER: 38,891
(C) REFERENCE/DOCKET NUMBER: P31353

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: 610-270-4478
- (B) TELEFAX: 610-270-5090
- (C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1974 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATGGCTAAAG	AAACATTTTA	TATAACAACC	CCAATATACT	ATCCTAGTGG	GAATTTACAT	60
ATAGGACATG	CATATTCTAC	AGTGGCTGGA	GATGTTATTG	CAAGATATAA	GAGAATGCAA	120
GGATATGATG	TTCGTTATTT	GACTGGAACG	GATGAACACG	GTCAAAAAAT	TCAAGAAAAA	180
GCTCAAAAAG	CTGGTAAGAC	AGAAATTGAA	TATTTGGATG	AGATGATTGC	TGGAATTAAA	240
CAATTGTGGG	CTAAGCTTGA	AATTTCAAAT	GATGATTTTA	TCAGAACAAC	TGAAGAACGT	300
CATAAACATG	TCGTTGAGCA	AGTGTTTGAA	CGTTTATTAA	AGCAAGGTGA	TATCTATTTA	360
GGTGAATATG	AAGGTGCGTA	TTCTGTTCCG	GATGAAACAT	ACTATACAGA	GTCACAATTA	420
GTAGACCCAC	AATACGAAAA	CGGTAAAATT	ATTGGTGGCA	AAAGTCCAGA	TTCTGGACAC	480
GAAGTTGAAC	TAGTTAAAGA	AGAAAGTTAT	TTCTTTAATA	TTAGTAAATA	TACAGACCGT	540
TTATTAGAGT	TCTATGACCA	AAATCCAGAT	TTTATACAAC	CACCATCAAG	AAAAAATGAA	600
ATGATTAACA	ACTTCATTAA	ACCAGGACTT	GCTGATTTAG	CTGTTTCTCG	TACATCATTT	660
AACTGGGGTG	TCCATGTTCC	GTCTAATCCA	AAACATGTTG	TTTATGTTTG	GATTGATGCG	720
TTAGTTAACT	ATATTTTCAGC	ATTAGGCTAT	TTATCAGATG	ATGAGTCACT	ATTTAACAAA	780
TACTGGCCAG	CAGATATTCA	TTTAATGGCT	AAGGAAATTG	TGCGATTCCA	CTCAATTATT	840
TGGCCTATTT	TATTGATGGC	ATTAGACTTA	CCGTTACCTA	AAAAAGTCTT	TGCACATGGT	900
TGGATTTTGA	TGAAAGATGG	AAAAATGAGT	AAATCTAAAG	GTAATGTTGT	AGACCCTAAT	960
ATTTTAATTG	ATCGCTATGG	TTTAGATGCT	ACACGTTATT	ATCTAATGCG	TGAATTACCA	1020
TTTGGTTCAG	ATGGCGTATT	TACACCTGAA	GCATTTGTTG	AGCGTACAAA	TTTCGATCTA	1080
GCAAATGACT	TAGGTAACCT	AGTAAACCGT	ACGATTTCTA	TGGTTAATAA	GTACTTTGAT	1140
GGCGAATTAC	CAGCGTATCA	AGGTCCACTT	CATGAATTAG	ATGAAGAAAT	GGAAGCTATG	1200
GCTTTAGAAA	CAGTGAAAAG	CTACACTGAA	AGCATGGAAA	GTTTGCAATT	TTCTGTGGCA	1260
TTATCTACGG	TATGGAAGTT	TATAAGTAGA	ACGAATAAGT	ATATTGACGA	AACAACGCCT	1320
TGGGTATTAG	CTAAGGACGA	TAGCCAAAAA	GATATGTTAG	GCAATGTAAT	GGCTCACTTA	1380
GTTGAAAATA	TTTCGTTATG	AGCTGTATTA	TTACGTCCAT	TCTTAACACA	TGCGCCGAAA	1440
GAGATTTTGT	AACAATTGAA	CATAACAAT	CCTCAATTTA	TGGAATTTAG	TAGTTTAGAG	1500
CAATATGGTG	TGCTTACTGA	GTCAATTATG	GTTACTGGGC	AACCTAAACC	TATTTTCCCA	1560

```

AGATTGGATA GCGAAGCGGA AATTGCATAT ATCAAAGAAT CAATGCAACC GCCTGCTACT 1620
GAAGAGGAAA AAGAAGAGAT TCCTAGCAAA CCTCAAATTG ATATTAAAGA CTTTGATAAA 1680
GTTGAAATTA AGGCAGCAAC GATTATTGAT GCTGAACATG TTAAGAAGTC AGATAAGCTT 1740
TTAAAAATTC AAGTAGACTT AGATTCTGAA CAAAGACAAA TTGTATCAGG AATTGCCAAA 1800
TTCTATACAC CAGATGATAT TATTGGTAAA AAAGTAGCAG TTGTTACTAA CCTGAAACCA 1860
GCTAAATTAA TGGGACAAAA ATCTGAAGGT ATGATATTAT CTGCTGAAAA AGATGGTGTA 1920
TTAACCTTAG TAAGTTTACC AAGTGCAATT CCAAATGGTG CAGTGATTAA ATAA 1974

```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 657 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

Met Ala Lys Glu Thr Phe Tyr Ile Thr Thr Pro Ile Tyr Tyr Pro Ser
 1             5             10             15
Gly Asn Leu His Ile Gly His Ala Tyr Ser Thr Val Ala Gly Asp Val
      20             25             30
Ile Ala Arg Tyr Lys Arg Met Gln Gly Tyr Asp Val Arg Tyr Leu Thr
      35             40             45
Gly Thr Asp Glu His Gly Gln Lys Ile Gln Glu Lys Ala Gln Lys Ala
      50             55             60
Gly Lys Thr Glu Ile Glu Tyr Leu Asp Glu Met Ile Ala Gly Ile Lys
      65             70             75             80
Gln Leu Trp Ala Lys Leu Glu Ile Ser Asn Asp Asp Phe Ile Arg Thr
      85             90             95
Thr Glu Glu Arg His Lys His Val Val Glu Gln Val Phe Glu Arg Leu
      100            105            110
Leu Lys Gln Gly Asp Ile Tyr Leu Gly Glu Tyr Glu Gly Trp Tyr Ser
      115            120            125
Val Pro Asp Glu Thr Tyr Tyr Thr Glu Ser Gln Leu Val Asp Pro Gln
      130            135            140
Tyr Glu Asn Gly Lys Ile Ile Gly Gly Lys Ser Pro Asp Ser Gly His
      145            150            155            160
Glu Val Glu Leu Val Lys Glu Glu Ser Tyr Phe Phe Asn Ile Ser Lys
      165            170            175
Tyr Thr Asp Arg Leu Leu Glu Phe Tyr Asp Gln Asn Pro Asp Phe Ile
      180            185            190

```

Gln	Pro	Pro	Ser	Arg	Lys	Asn	Glu	Met	Ile	Asn	Asn	Phe	Ile	Lys	Pro	195	200	205
Gly	Leu	Ala	Asp	Leu	Ala	Val	Ser	Arg	Thr	Ser	Phe	Asn	Trp	Gly	Val	210	215	220
His	Val	Pro	Ser	Asn	Pro	Lys	His	Val	Val	Tyr	Val	Trp	Ile	Asp	Ala	225	230	235
Leu	Val	Asn	Tyr	Ile	Ser	Ala	Leu	Gly	Tyr	Leu	Ser	Asp	Asp	Glu	Ser	245	250	255
Leu	Phe	Asn	Lys	Tyr	Trp	Pro	Ala	Asp	Ile	His	Leu	Met	Ala	Lys	Glu	260	265	270
Ile	Val	Arg	Phe	His	Ser	Ile	Ile	Trp	Pro	Ile	Leu	Leu	Met	Ala	Leu	275	280	285
Asp	Leu	Pro	Leu	Pro	Lys	Lys	Val	Phe	Ala	His	Gly	Trp	Ile	Leu	Met	290	295	300
Lys	Asp	Gly	Lys	Met	Ser	Lys	Ser	Lys	Gly	Asn	Val	Val	Asp	Pro	Asn	305	310	315
Ile	Leu	Ile	Asp	Arg	Tyr	Gly	Leu	Asp	Ala	Thr	Arg	Tyr	Tyr	Leu	Met	325	330	335
Arg	Glu	Leu	Pro	Phe	Gly	Ser	Asp	Gly	Val	Phe	Thr	Pro	Glu	Ala	Phe	340	345	350
Val	Glu	Arg	Thr	Asn	Phe	Asp	Leu	Ala	Asn	Asp	Leu	Gly	Asn	Leu	Val	355	360	365
Asn	Arg	Thr	Ile	Ser	Met	Val	Asn	Lys	Tyr	Phe	Asp	Gly	Glu	Leu	Pro	370	375	380
Ala	Tyr	Gln	Gly	Pro	Leu	His	Glu	Leu	Asp	Glu	Glu	Met	Glu	Ala	Met	385	390	395
Ala	Leu	Glu	Thr	Val	Lys	Ser	Tyr	Thr	Glu	Ser	Met	Glu	Ser	Leu	Gln	405	410	415
Phe	Ser	Val	Ala	Leu	Ser	Thr	Val	Trp	Lys	Phe	Ile	Ser	Arg	Thr	Asn	420	425	430
Lys	Tyr	Ile	Asp	Glu	Thr	Thr	Pro	Trp	Val	Leu	Ala	Lys	Asp	Asp	Ser	435	440	445
Gln	Lys	Asp	Met	Leu	Gly	Asn	Val	Met	Ala	His	Leu	Val	Glu	Asn	Ile	450	455	460
Arg	Tyr	Ala	Ala	Val	Leu	Leu	Arg	Pro	Phe	Leu	Thr	His	Ala	Pro	Lys	465	470	475
Glu	Ile	Phe	Glu	Gln	Leu	Asn	Ile	Asn	Asn	Pro	Gln	Phe	Met	Glu	Phe	485	490	495
Ser	Ser	Leu	Glu	Gln	Tyr	Gly	Val	Leu	Thr	Glu	Ser	Ile	Met	Val	Thr	500	505	510
Gly	Gln	Pro	Lys	Pro	Ile	Phe	Pro	Arg	Leu	Asp	Ser	Glu	Ala	Glu	Ile	515	520	525
Ala	Tyr	Ile	Lys	Glu	Ser	Met	Gln	Pro	Pro	Ala	Thr	Glu	Glu	Glu	Lys	530	535	540

37